



Pipelines de ETL e Machine Learning com Apache Spark

O Que é Processamento Distribuído?

O processamento distribuído refere-se a um sistema em que múltiplas máquinas ou computadores trabalham juntas para realizar uma tarefa ou alcançar um objetivo comum. Nesse sistema, a carga de trabalho é dividida entre os computadores, permitindo que eles processem informações simultaneamente e, muitas vezes, de forma independente. O processamento distribuído é uma subcategoria da computação distribuída, que engloba uma ampla variedade de sistemas que distribuem tarefas entre múltiplos computadores.

Citamos a seguir algumas características e benefícios do processamento distribuído:

- **Descentralização:**

Em vez de depender de um único computador central, o processamento distribuído utiliza vários computadores ou nós. Isso pode aumentar a resiliência do sistema, pois a falha de um único nó geralmente não resulta na falha de todo o sistema.

- **Escalabilidade:**

É possível adicionar mais nós ao sistema para lidar com cargas de trabalho maiores ou para melhorar o desempenho.

- **Paralelismo:**

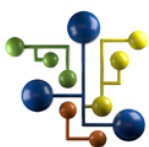
Múltiplas operações ou tarefas podem ser executadas simultaneamente, aproveitando o poder de processamento de vários nós.

- **Flexibilidade:**

Os sistemas de processamento distribuído podem ser configurados para atender a diferentes necessidades e podem ser ajustados conforme as demandas mudam.

- **Eficiência:**

Ao distribuir tarefas entre vários nós, é possível otimizar o uso de recursos e acelerar o processamento.



Equipe DSA

Muito Obrigado!
Continue Trilhando Uma Excelente Jornada de Aprendizagem.